Does smartphone usage behavior influence willingness to agree to passive data collection via a mobile app?

Jnlenth Martinez Atencio, University of Neuchâtel
Caroline Roberts, University Lausanne
Jessica Herzing, University of Bern
Daniel Gatica-Perez, EPFL
Mobile Device Use in Switzerland

- Smartphone market penetration of 73%
  Bankmycell.com, 2020

- On average Swiss people own 3.4 connected devices
  Statista.com, 2020

- 85% prefer a mobile device for going online
  Koptyug, 2020
“...simply possessing a mobile device does not necessarily indicate a willingness to use it for mobile responding”

(de Bruijne et al., 2014, p. 731)
Willingness to Participate in Passive Data Collection

1. General reasons which influence survey participation
   - Respondent's interest in the topic
   - Questionnaire length/ Demanded time for participation

2. Level of data apprehension
   - Attitudes towards their data privacy

3. Familiarity with new technologies
   - Technology skills' level
   - Data collection knowledge

Keusch et al., 2019
What could influence willingness?

Willingness to agree to passive data collection

- Different mobile technologies
  
  (Jäckle et al., 2019; Wenz et al., 2019)

- Incentives, topic of study, or even skill level of smartphone use
  
  (Keusch et al., 2019)

- Focus of quantity vs type of activities
  
  (Keusch et al., 2021; Struminskaya et al., 2021)
Research Questions

• RQ1: Can smartphone users be differentiated in terms of the types of activities that they use their devices for?
• RQ2: What explains variation in smartphone usage behaviours (e.g., socio demographic characteristics, attitudes)?
• RQ3: How do smartphone usage habits relate to willingness to agree to passive data collection via an app?
Sample size consisted of 2175 individuals from the French-Speaking part of Switzerland.

Experimental design:
- Group 1 (n=1’088) – online panel via a web browser
- Group 2 (n=1’087) – online panel via mobile application

31.6% (n=687) response rate for the first wave in both methods of data collection.

Dataset was filtered to select only those who had a smartphone (n=570).

Roberts et al., 2020
RQ1: Can smartphone users be differentiated in terms of the types of activities that they use their devices for?
Analytical Approach – Exploratory Methods

Cluster Analysis:
- Agglomerative method
- Behavior variables (i.e., activities declared to do with their smartphones)

Multiple Correspondence Analysis MCA:
- 3 different scenarios for active variables: Behavior, Attitudes, Combination
- 10 top contributors
Results: Cluster Analysis

Cluster 1: Data – Risk Tolerant

Cluster 2: Data – Risk Averse

N=371

N=183
Results: Cluster Analysis

Cluster 1: Data – Risk Tolerant

Cluster 2: Data – Risk Averse

N=371

N=183
### Results: Cluster Analysis

#### Cluster 1: Data – Risk Tolerant

<table>
<thead>
<tr>
<th>Activity</th>
<th>Cluster One: Data – Risk Tolerant</th>
<th>Cluster Two: Data – Risk Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Browsing online</td>
<td>96.5%</td>
<td>91.3%</td>
</tr>
<tr>
<td>Reading/extending emails</td>
<td>93.9%</td>
<td>83.6%</td>
</tr>
<tr>
<td>Taking photographs</td>
<td>99.5%</td>
<td>92.9%</td>
</tr>
<tr>
<td>Using social media platforms</td>
<td>91.6%</td>
<td>89.1%</td>
</tr>
<tr>
<td>Posting on social media platforms</td>
<td>74.9%</td>
<td>99.5%</td>
</tr>
<tr>
<td>Shopping</td>
<td>72.5%</td>
<td>41.0%</td>
</tr>
<tr>
<td>Banking</td>
<td>56.0%</td>
<td>25.7%</td>
</tr>
<tr>
<td>Installing new apps</td>
<td>86.0%</td>
<td>59.0%</td>
</tr>
<tr>
<td>Using GPS apps</td>
<td>88.3%</td>
<td>73.3%</td>
</tr>
<tr>
<td>Connecting to other devices via Bluetooth</td>
<td>54.4%</td>
<td>42.1%</td>
</tr>
<tr>
<td>Gaming</td>
<td>52.3%</td>
<td>47.3%</td>
</tr>
<tr>
<td>Media</td>
<td>90.6%</td>
<td>59.0%</td>
</tr>
</tbody>
</table>

N=371

#### Cluster 2: Data – Risk Averse

<table>
<thead>
<tr>
<th>Activity</th>
<th>Cluster One: Data – Risk Tolerant</th>
<th>Cluster Two: Data – Risk Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Browsing online</td>
<td>96.5%</td>
<td>91.3%</td>
</tr>
<tr>
<td>Reading/extending emails</td>
<td>93.9%</td>
<td>83.6%</td>
</tr>
<tr>
<td>Taking photographs</td>
<td>99.5%</td>
<td>92.9%</td>
</tr>
<tr>
<td>Using social media platforms</td>
<td>91.6%</td>
<td>89.1%</td>
</tr>
<tr>
<td>Posting on social media platforms</td>
<td>74.9%</td>
<td>99.5%</td>
</tr>
<tr>
<td>Shopping</td>
<td>72.5%</td>
<td>41.0%</td>
</tr>
<tr>
<td>Banking</td>
<td>56.0%</td>
<td>25.7%</td>
</tr>
<tr>
<td>Installing new apps</td>
<td>86.0%</td>
<td>59.0%</td>
</tr>
<tr>
<td>Using GPS apps</td>
<td>88.3%</td>
<td>73.3%</td>
</tr>
<tr>
<td>Connecting to other devices via Bluetooth</td>
<td>54.4%</td>
<td>42.1%</td>
</tr>
<tr>
<td>Gaming</td>
<td>52.3%</td>
<td>47.3%</td>
</tr>
<tr>
<td>Media</td>
<td>90.6%</td>
<td>59.0%</td>
</tr>
</tbody>
</table>

N=183
Results: Multiple Correspondence Analysis

**Behavior**

- Contrib. of variables to Dim. 1
- Contrib. of variables to Dim. 2

**Attitudes**

- Contrib. of variables to Dim. 1
- Contrib. of variables to Dim. 2

**Combination**

- Contrib. of variables to Dim. 1
- Contrib. of variables to Dim. 2
Graphical visualization of the combination
Graphical visualization of the combination

Dark blue ellipse represents first dimension
Graphical visualization of the combination

Dark blue ellipse represents the first dimension

Green ellipse represents the second dimension
Smartphone users can be differentiated in terms of the type of activities that they use their devices for.
RQ2: What explains variation in smartphone usage behaviors (e.g., socio demographic characteristics, attitudes)?
Analytical Approach – Inferential Statistics

- Personal data being collected
- Data shared without consent
- Identity theft
- Data used for personalized ads

- Likes technology
- The internet as a communication tool
- Problem solving knowledge
- Being anonymous
- The internet as a privacy threat
- The internet is trustworthy

- Gender
- Age category
- Education level
- Frequency of smartphone use

Prob. of belonging to Cluster 1 = Attitudes Variables – Data + Attitudes Variables – Internet

Socio demographic
Results: Logistic Regression

- Female: Exp(B) = 1.57*
- ≤40 years old: Exp(B) = 2.82***
- High smartphone frequency use: Exp(B) = 2.99**
- A little bit worried of personal data being collected: Exp(B) = 7.44**

*p<0.05; **p<0.01; ***p<0.0001
Variation is explained by sociodemographic characteristics

- 86% of people between 18-29 y/o and 77% between 30-49 y/o are users
- 54% of users are female
- On average users spend 35 minutes a day on the platform. 96% through a smartphone

- 67% of people between 18-29 y/o are users
- 51% are females
- On average users spend 53 minutes a day on the platform
RQ3: How do smartphone usage habits relate to willingness to agree to passive data collection tasks?
Analytical Approach – Inferential Statistics

Total Sum Score of Willingness to Agree to Passive Data Collection = Sum Score of Attitudes towards their data + Sum Score of Attitudes towards the internet + Cluster 1 belonging + Socio demographic

- Range between 3 to 12 points
- Range between 4 to 20 points
- Range between 6 to 30 points
- Dummy variable:
  - 1 = Belongs to cluster 1
  - 0 = Does not belong to cluster 1
- Gender
- Age category
- Education level
- Frequency of smartphone use
Results: OLS Regression

≤ 40 years old
B = 0.56***

Sum Scores - Attitudes towards the internet
B = 0.11***

Sum Scores - Attitudes towards their data being online
B = 0.18***

Data Risk – Tolerant (Cluster 1)
B = 0.39.
Willingness is higher when:

- Respondents are young
- Positive attitudes towards the internet
- Positive attitudes towards their data being online
- Data exposition is a normal behavior

(Revilla et al., 2019; Wenz et al., 2019; Keush et al., 2019; Jäckle et al., 2019; Anshari et al., 2016).

Smartphone usage habits with apps where data is exposed are positively related to willingness.
Limitations

From survey
- Specific target population, low response rates
- Different methods of data collection; samples pooled

From analysis
- Only stated willingness was measured
- Both experimental groups were considered as one
Conclusion

• Smartphone users can be differentiated by the activities that they use their smartphones for
• This differentiation is mainly explained by sociodemographic characteristics
• There is a connection between using applications where data is highly exposed and stated willingness to agree to passive data collection
• Attitudes seem to be more important to explain hypothetical willingness, yet it will be important to see how behaviors might be linked to actual compliance of requests.
• Results confirm: Keush et al. (2019); Jäckle et al. (2019); Revilla et al. (2019); Wenz et al. (2019); Anshari et al. (2016); Tessem et al. (2019)
THANK YOU FOR YOUR ATTENTION
ACKNOWLEDGEMENTS

• Special thanks to:
  • Philip Abbet (Idiap Research Institute)
  • Jimena Sobrino Piazza (UNIL)
  • Stéphane Bonny (UNIL)
  • Maud Reveilhac (UNIL)
  • Data Collection and Analysis and Political Surveys teams at FORS

• CROSS Funding Program – EPFL-UNIL
• FORS: Additional funding for incentives and printing costs for the field experiment, as well as technical support
• SSP/ISS: Additional funding for research assistants